AD_____

GRANT NUMBER: DAMD17-94-J-4387

TITLE: Adding Data Accessibility and Rule-Based Targeted Data Collection to the California Cancer Reporting System for Breast Cases

PRINCIPAL INVESTIGATOR: Doctor Barry Gordon

CONTRACTING ORGANIZATION: California Public Health Foundation
Berkeley, California 94704-1103

REPORT DATE: September 1996

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, MD 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

DTIC QUALITY INSPECTED 1

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1996	Annual (15 Aug 95 - 14 Aug 96)

4. TITLE AND SUBTITLE
Adding Data Accessibility and Rule-Based Targeted Data Collection to the California Cancer Reporting System for Breast Cases

5. FUNDING NUMBERS
DAMD17-94-J-4387

6. AUTHOR(S)
Doctor Barry Gordon

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
California Public Health Foundation
Berkeley, California 94704-1103

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick
Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

19970113 056

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We have completed the second year of our three-year contract. As a major deliverable, we have opened public access to California cancer case information, allowing anyone on the World Wide Web to propose data queries and receive tables of counts from their query. Information available on our askcnet.org web site include:

- Cancer types, stages, and treatment for all California cancer cases diagnosed 1988-1993 (N=780,804)
- Preliminary results of 1995 study fields on cancer screening, protocol use, and clinical indicators
- SEER's 10% of national cancer cases 1973-1992 (N~2 million) (awaiting final signoff by SEER)

Coming soon to the site will be a clinical trials matching system, allowing the general public and healthcare professionals to obtain lists and consumer-oriented summaries of current IRB-approved clinical trials that match requested patient characteristics.

We are in the process of implementing a combined web and BBS-based telecommunications solution for connecting California hospitals, the eight Regional Registries, and our central office. We are stationing NT-based Wildcat communications servers at each region This allows email and data transmittal to occur either through dialup connections or, with the proper security options, through the Internet.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer, web, query, Internet, registry, SEER, clinical trials	32
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet *optical scanning requirements*.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	PR	- Project
G	- Grant	TA	- Task
PE	- Program Element	WU	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- DOE - See authorities.
- NASA - See Handbook NHB 2200.2.
- NTIS - Leave blank.

Block 12b. Distribution Code.

- DOD - Leave blank.
- DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- NASA - Leave blank.
- NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

BJ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

BC For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

9/12/96

PI - Signature Date

Table of Contents

Introduction

Background

The California Cancer reporting system has been collecting data on all breast cancer cases (~20,000/year) for over eight years statewide, and for more than 25 years in some geographic areas. This pioneering system is built around mandatory electronic cancer case reporting, coupled with software specifically designed to carry out this mandatory reporting.

Hospital-based tumor registry software (C/NET) was developed by the California Public Health Foundation (CPHF) to achieve high-quality and efficient data collection. CPHF has distributed C/NET to 250+ hospitals and other cancer reporting facilities throughout the state. C/NET is also used on portable computers to collect cases from the remaining, smaller hospitals around the state. C/NET currently handles autocoding, data entry, interfield edits, and the preparation of periodic transmit files for conveying new case reports and later case modifications or corrections to California's central cancer registry. Cases are first sent to one of eight regional intrastate registries, which carry out further editing, case matching, and quality control before sending the cases on to the state central registry.

The C/NET hospital software is widely used and appreciated, as it carries out functions essential to hospital registries and their clinician staffs beyond case collection and reporting. Various versions of C/NET are in use in over 650 other hospitals outside of California. These versions have been distributed by the American College of Surgeons (ACoS), and, more recently, by the California Public Health Foundation, our parent organization. C/NET software gathers data on cancer diagnoses, stage, treatment, and followup in national standard format. It includes all fields necessary for the National Cancer Institute (NCI)'s SEER Program and ACoS's National Cancer Database. However, not all data items gathered at the hospitals have been uploaded to the California central registry at present.

As the number of years of complete statewide data collection has increased, the California Cancer Registry has used its statewide database in numerous studies and publications.

Statement of the Problem

Several factors have hindered wide usage of registry data for clinical and epidemiological studies in breast cancer. These problems include: lack of easy access to the data, built-in delays in the reporting system, difficulty in pilot testing or adding new data fields targeted for breast cancer studies, and lack of rapid case ascertainment to support special studies. Many investigators have noted the imbalance in registry focus, with more attention being given to data collection than data usage. Also, analytical tables and reports are not

available through a single, uniform access method. Instead they are scattered across many locally produced publications. This has limited access to statewide cancer registry data to a few sophisticated users, who are largely within the cancer reporting system.

Most population-based registries must wait until complete years are closed out, including a time-consuming death clearance follow-back, before data can be shown. The delay is often two years or more. While the time required to produce accurate cancer rates is difficult to shorten, faster pathways for reviewing preliminary data are needed. Breast cancers, in particular, may be identifiable to a large degree through fast channels such as path reports.

Interview studies are particularly sensitive to the timeliness of case identification. Because registry reporting is too slow at present, most special studies must mount a labor-intensive separate casefinding effort which duplicates the existing system. A unified rapid reporting system is needed.

The current cancer reporting systems are also slow in their ability to revise data collection to match changing clinical practices. Potential new methods of detection, staging, and treatment are constantly appearing and requiring evaluation. Although computers should facilitate the rapid deployment of new fields and codes for pilot testing, this is not occurring. Rapid-turnaround studies require both faster methods for distributing data item revisions, as well as faster methods for reviewing early results.

Purpose and Technical Objectives

We have been adding new functions to the computerized reporting system now in place in order to solve these problems. We want to make the data available in highly useable form and foster studies not before possible on such a large and ethnically diverse population. Examples include testing whether access to care for breast cancer and quality of care are related to patient insurance status, ethnic, type of provider, or geographic factors.

Specific goals of the system enhancements include:

1. Provide a **dial-in user query system** open to all investigators, allowing user-defined tabulations, plots, and offline maps on a wide variety of breast case data. Databases open to queries would come from three sources: statewide data from 1988 to the current year, NCI SEER data from 1973 forward, and data from other states, by their agreement, who have submitted to the North American Association of Central Cancer Registries (NAACCR).

2. Facilitate the rapid, electronic communication of case-related information by adding **integrated telecommunications** to the C/NET hospital-based registry software. This would allow all 250 hospitals in California to easily submit case data, receive shared data, query the central statistical databases, carry out e-mail communication with their

peers and their technical support personnel, and load comparison data counts which parallel their own in-house queries, in order to produce tables and graphs with direct comparisons.

3. Encourage cost-effective targeted patient interview studies by adding a flexible **rule-based rapid case-reporting** channel for cases discovered in hospitals that potentially meet special patient interview study criteria, so the regional registries can conduct these studies at a much reduced cost.

4. **Monitor and encourage enrollment in treatment protocols** by adding rule-based criteria to hospital registry software. The rules would flag cases meeting current criteria for one or more national treatment protocols. This helps track the rate of use of these protocols in the target patient population, and may facilitate the recruitment of patients.

5. Help **solve at the hospital level the problem of incomplete treatment reporting** by adding Physician Data Query (PDQ) recommended treatment plans to the software. This will identify the most commonly recommended treatment for patients of a particular age, stage, and disease type. The software will also support the automatic mailing of physician letters to inquire whether this therapy has been carried out, if the medical record is incomplete.

6. Assure that a large cohort of breast cases has crucial staging and followup information by carrying out a one-time catch-up data upload of **AJCC stage and followup data** from California hospitals to the regional and central registry. These data are mostly available in the hospital systems, but have not been fully communicated to the central registry. They will be transmitted on a regular basis thereafter.

7. Facilitate the timely flow of case-related data by setting up **regional registry bulletin board systems** in California to handle all communications with hospitals and other treatment facilities in eight regional centers. These bulletin board systems (BBS's) would manage automatic information uploads and downloads, as well as e-mail, and would be tied together into a statewide network. They can also be used to update data collection software.

8. Broaden the kinds of research questions that can be studied statewide by **adding new fields** to the recommended state data set, and adding rapid-turnaround study fields that would change every year. To study protocol enrollment, protocol number or reason for no protocol would be coded. To study financial limitations on care for breast patients, insurance status/source of payment would be tracked. In order to study treatment success, disease free interval would be calculated. Fields will also be added to collect co-morbid conditions, and to pass on the clinical indicators for breast cancer required by the JCAHO beginning in 1995. To encourage hospital studies in more

detail, optional fields would include procedures performed and their costs, and detail on radiation and chemotherapy performed or the reason none was given. These fields would be evaluated and revised yearly.

Work Accomplished

The following list corresponds to the first and second year workplans submitted in the original proposal.

1. **Set up bulletin board hardware and LAN for the central registry node, and install BBS software**. Completed in year one.

2. **Add proposed new data items to C/NET hospital registry software and distribute to all registry hospitals in California.** Completed in year one.

3. **Announce new query bulletin board, provide user documentation, and demonstrate its use to groups of potential users.** Completed in year one.

4. **Install the HIRS query system, and load California registry data.** Completed in year one.

5. **Implement e-mail and case transmit bulletin boards in 7 regional registries.**

Two California regions have been running BBS systems, in addition to our central C/NET BBS system. During the process of testing email sharing between these systems, we became concerned that once the current BBS's were deployed in all of the regions, they would quickly become technically outdated. Our primary concern was that the existing systems were using DOS-based software and dial-up modem lines. The computer industry has been phasing out DOS-based programs in favor of Windows-based solutions. Additionally, the Internet has been gaining widespread acceptance. The new communications model is one in which users connect to an Internet service provider close to their location and access information providers over a shared network, rather than the older model of directly dialing a specific information provider using a modem and phone line.

Since that time, Mustang, the developer of our DOS-based BBS software, announced its new BBS software designed to run on Windows-based computers. The software also addressed the concerns from a user standpoint. Users could access the new BBS using either existing DOS-based terminal programs or with a new Windows-based graphical program included with the BBS (see Appendix A). The new BBS also provides the ability for users to access the system via the Internet (using either text or graphical software) as well as traditional modem dial-up lines.

The decision was made to deploy the new version of the BBS software to the regions. We had to wait until the new Internet-aware software was available, causing some delay in the deployment of the new systems. Due to the additional hardware requirements (a faster

processor and more RAM) of the new software, the decision was made to also supply BBS hardware to the regions. Currently, the first four computer systems have been purchased. Two of these systems will be deployed with the new software in several weeks, quickly followed by the next two systems. Once we have resolved any initial problems with these systems, the remaining regions will be brought online.

See Appendix A for examples of the BBS screens as viewed on the Internet and through a standard dial-up BBS connection.

6. **Carry out one-time upload of patient followup and AJCC staging data.**

This could not yet be carried out, since the regional systems still have not yet been modified to receive the data. Unfortunately, these systems are not under our direct control. The central and regional staffs now realize the importance of this step, and it should be accomplished by the end of 1996.

7. **Load 1993 case data into query system when available.**

The public use file from the California Cancer Registry (CCR) of incidence cases diagnosed between 1988 and 1993 was received in April, 1996. This file was converted into a format appropriate for loading into the query database, and the interface between the database and the World Wide Web was created. After negotiations with the CCR concerning the protection of confidentiality and the monitoring of the level of usage, as well as requesting additional data fields not available in the public use file, we created several web pages. One captures visitor characteristics, and another screen helps encourage user compliance with CCR requirements. For reporting to the CCR, we developed a database to capture the query requests that each visitor would make. With these features in place, we were able to make the query engine publicly accessible with CCR data in August, 1996.

See Appendix B for sample usage agreement, query request, and query output screens using California data.

8. **Load NCI SEER data into HIRS.**

We obtained the SEER CD-ROM version of their 1973-1992 database and converted and loaded it into the query engine. We are continuing negotiations with SEER on making their data publicly accessible via our web site. We anticipate receiving their approval after the SEER Principal Investigators meeting in October, 1996. Once we have their approval, we can make the SEER datasets available very quickly. We have provided SEER with a draft screen for protection of confidentiality (see Appendix D) and a summary of the visitors' report, based on the system we created for the California Cancer Registry public use file. These may be modified quite easily to the specifications provided by SEER. SEER has been releasing summary tables of data through 1993 at its own website, so we expect a public use

file with 1993 data to be made available soon. See Appendix C for an example data query of SEER breast cancer cases 1973-1992, tabulating surgery versus radiation given.

One feature that NCI's SEER program requested was a report for each query run on their data, showing the requester and the exact query performed. We have designed a system to accomplish this detailed logging, and are using it to log and report each of our California data requests as well. This useful log provides us detailed feedback about what kinds of questions are being posed, and by type of user.

See Appendix E for an example of a report of several query sessions run recently on the Web site.

9. Add 1994 cases from state and SEER when available.

Complete data on California 1994 cases will be available in March 1997, and we will include the data in our query database as soon it is officially announced. The SEER data for 1994 cases will not be available until approximately August 1997.

10. Prepare new query system database of hospital breast case reports, 1988-94, including preliminary data on the new fields added in year one.

In January, 1996, we began collecting sample files from selected California regional registries of 1995 cases with the new data fields added for 1995 cases. The early files showed a low rate of completion of the new fields by registrars. We hoped that in time, the utilization of these new fields would increase. A second round of sample files, representing about 25% of the year's cases and 168 participating hospitals, showed some improvement.

One complication that delayed our receiving cases was that the regional registries underwent a database conversion between May and July. The ability to create the files we required was postponed until after the conversion was verified to be accurate. We requested more cases from the regional registries in August and have received files from only half of the registries.

We conducted a preliminary analysis of the files for levels of utilization of the new fields. We noted that some hospitals (approximately 30% of the participating hospitals) coded a reasonably high percentage of cases, while others (approximately 50% of the participating hospitals) utilized some fields but did not enter codes for others, and some hospitals did not code the fields at all. Overall there was still a very low level of utilization of the new data fields. We decided to select cases only from those hospitals with a consistent pattern of utilization, and we created analysis files that included only hospital records in which all of the fields within a group were coded.

See Appendix F for a list of the fields grouped. Appendix G illustrates the output using one of the better utilized fields, Discovered by Screening.

11. Revise new data items, depending on results of review of previous year's testing, and new research needs.

In an effort to decide on data revisions, the C/NET staff met with our C/NET advisors group, which includes cancer registrars serving as representatives from ten hospitals who have been appointed by the California regional registries. After discussion of the year one dataset, the following additions were made for year two:

Added Fields

Global clinical indicators:

- *Mammogram* (yes, no, unknown)
- *Barium Enema* (yes, no, unknown)
- *Ultrasound to Primary Site* (yes, no, unknown)

These were added to the list of clinical guidelines, since they are common tests that many hospitals consider part of a standard of care for certain cancer sites.

Tumor Markers:

- *CEA Value* (0-999.9)
- *CEA Level* (positive, negative, borderline, not done, unknown)

The advisors wanted to add this to the dataset to compare values of this marker, commonly used in colon cancer cases.

The following fields were added because of new requirements by the American College of Surgeons. They added these to the required or recommended dataset for hospital registries

- *First Name AKA*
- *Name Suffix*
- *Dates of Inpatient Admission & Discharge*
- New MD fields to identify *Radiation Oncologist & Medical Oncologist*
- *AJCC/TNM*: Separate fields for Clinical, Path, and Other coding of stage at diagnosis
- *TNM Coder*
- *Non-cancer-directed Summary/Hospital Surgery*
- *Date of Non-Cancer Directed Surgery*
- *Reconstructive Surgery Summary*
- *Reconstructive Surgery Hospital*
- *Pediatric Stage, Pediatric Staging System, Pediatric Stage Coder*

Deleted Fields

No comparison data items were chosen for removal from the data collection system at this time.

12. **Integrate telecommunications functions into C/NET.**

This will be accomplished after the new regional bulletin board systems are in place (see task 5 above). We have selected Wildcat Navigator as the client software that will handle telecommunications between hospitals and regional registries. Among the reasons for this selection are:

(a) Wildcat provided us with unlimited free copies when we purchased the Wildcat 5 BBS software for the regions.

(b) The software allows hospitals to use whichever communications connection is available to them, either dial up modem or in-house Internet LAN connection. If available, the Internet alternative would be free of telecommunications costs.

(c) The software allows us to either write scripts or program direct interaction between our C/NET registry and the communications software.

We will provide copies of Wildcat Navigator to all California hospital registries, contingent on their having a PC with Windows and either a modem or an Internet connection. The scripts and interface programming will be completed over the next year.

13. **Implement rapid case reporting in one region.**

This has not yet begun but will be implemented over the next year.

14. **Program rule-based recommended treatment system.**

We have completed the design of the rule-based recommended treatment system. Appendix H shows the rule to be implemented for breast cancer, based on NCI's PDQ. Implementing the rule in our 250 hospital cancer registry systems will be completed within the next year.

15. **Program rule-based flagging system to indicate when hospital registry cases may be eligible for existing treatment protocols.**

This is another area where changes in the technology since we submitted our proposal are allowing an improvement in the deployment methods. Our original intent was to provide updatable rules in the hospital software that would signal which cancer cases might be subjects for current treatment protocols. However, the task of keeping the hospital-based

rules system up-to-date in hundreds of locations was somewhat daunting. With the increased use of the Internet, another solution has become available. We decided to publish a clinical trials matching system as part of our Web site. This could more easily be kept up to date, and could provide links to NCI's trial descriptions, as well as to sites where patients are being recruited into trials. Thus cancer registrars, clinicians, and the general public could locate open trials that are relevant to their situation.

The Web site is in prototype right now. The matching is carried out by asking the inquirer to anonymously fill in certain data on their age group, type of cancer, stage, and demographics. They then see a list of the trials which match on these criteria. A patient can take this list to her or his physician for more detailed discussion. Cancer registrars can use the matching to check whether the case they are reporting might be suitable for a clinical trial.

Our current prototype has fifty active trials in the database. See Appendix I for an example of the Web form that is filled out to begin the match. Appendix J shows an example of the 'general public' summary of one particular trial. These summaries are provided by the Clinical Trials Information Program (CTIP), a community-based organization in California.

Our goal is to integrate the hospital-based desktop cancer registry application with web-based lookup of current information. One of the first uses will be to query our clinical trials web system during cancer case coding, in an effort to flag potential candidates for trials.

Conclusions

1. We have successfully implemented online Web-based queries using cancer case datasets of substantial size. For example, we loaded twenty years of data based on 10% of the US in a few hours. Queries against this database then take about one second to complete. Transmission of the results to the user depends on the speed of the Internet connection.

2. We find the detailed logging of incoming user queries to be feasible and useful in monitoring the kinds of uses the query site is being asked to provide. Users are willing to identify themselves by general category (e.g. researcher), making that information available for analysis as well.

3. It takes careful planning and cooperation with data suppliers in order to publish meaningful data queries on the web. Each agency, such as the California Cancer Registry and SEER, have their own concerns regarding case confidentiality, disclaimers, descriptions, and the monitoring of use. However, once we solved these problems for the CCR, it appears to be relatively easy to meet other agencies' similar requirements.

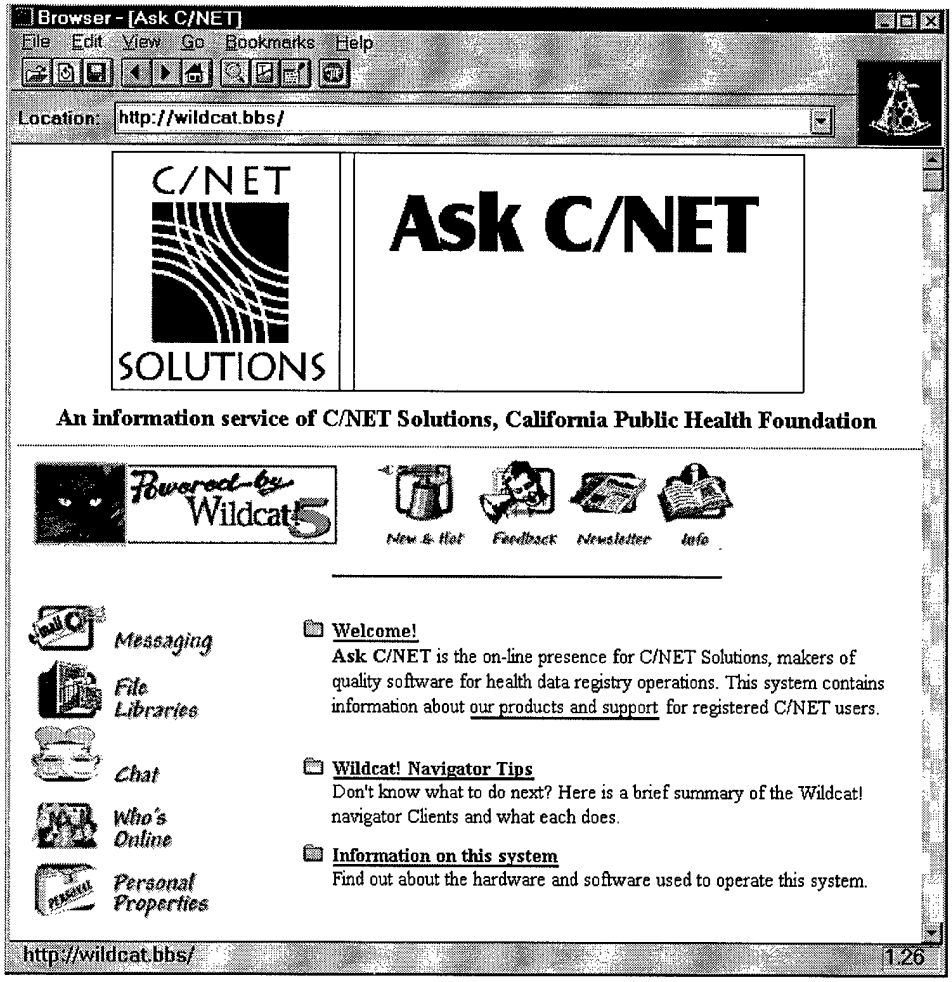
4. The Web appears to be a better delivery system for clinical trials information lookups than adding software to hospital registry systems. It gives the general public as well as the healthcare professional access to current trial selections that match specific criteria. Also, we find it easier to update the trials list and support links to other trials information sites through a central web server than through individual desktop software.

5. We are pleased that telecommunications options are now becoming more integrated. Bulletin board software suppliers such as Wildcat are now offering solutions that integrate the more traditional dialup model with the benefits of Internet connections and web browser-like screens. This makes it much easier for us to provide a single software package to our eight regions which can encompass the variety of user access methods. It also means we can develop a smooth interface between our hospital registry software and our Internet web content, such as clinical trials matching.

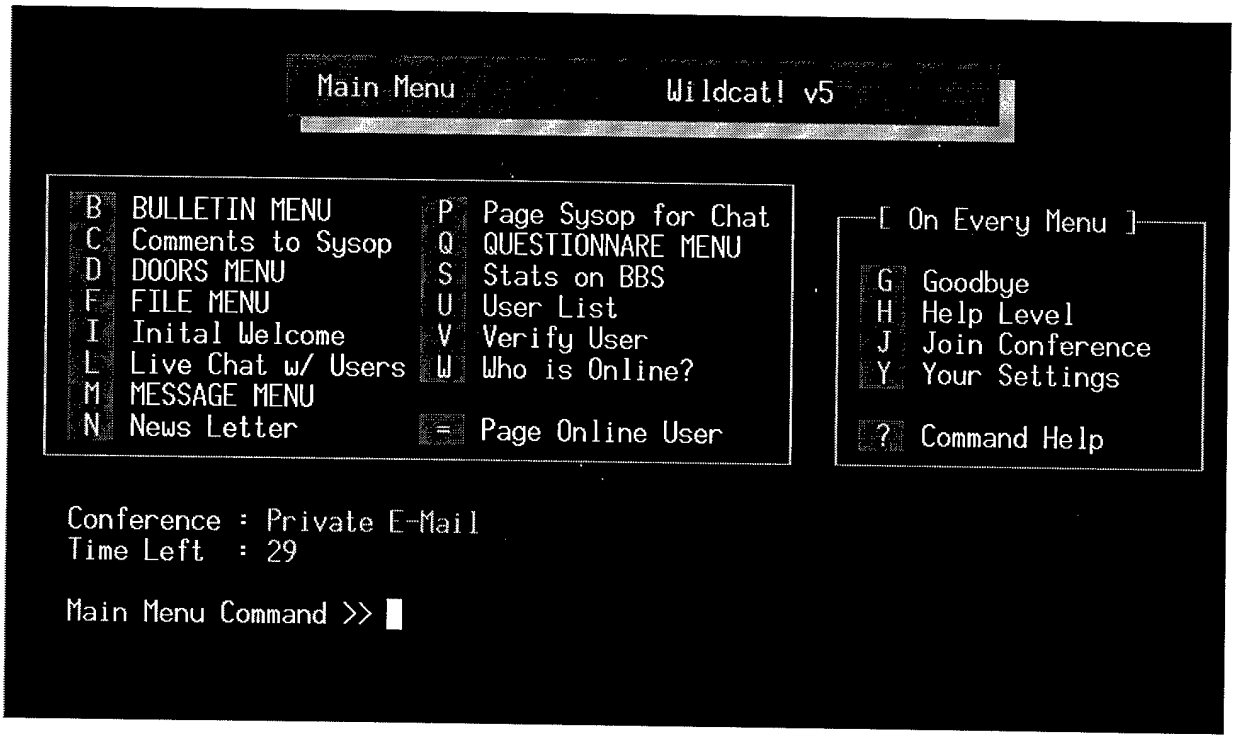
6. Rapid turnaround of registry data is still not easily accomplished. The major problem in California comes from the fact that a number of organizations are involved in the data flow, including regional registries with different software systems. These agencies have their own sets of priorities.

7. Implementing new optional data items takes more training resources than we thought. Each registrar must be convinced of the value of the extended dataset, and decentralized training meetings are needed to make the coding consistent and answer the inevitable questions that arise when using new data items.

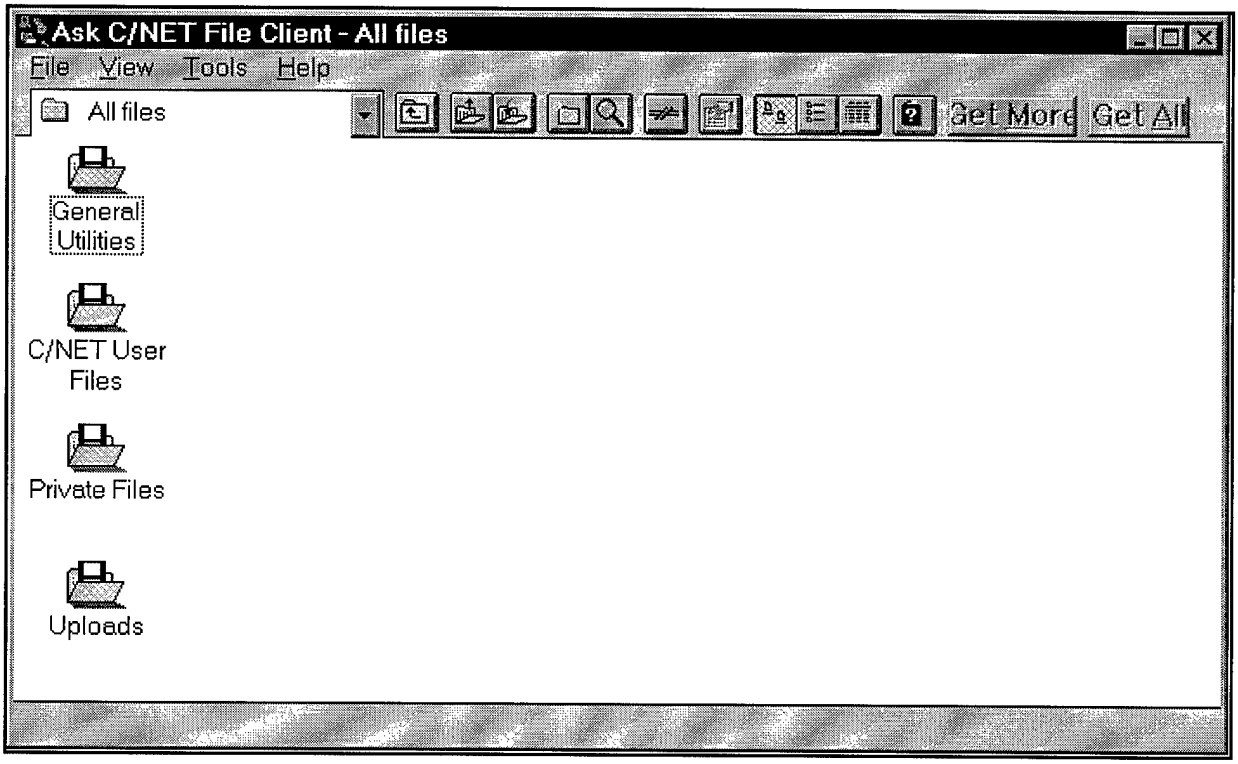
Appendix A-1
California Regional Bulletin Board Menus as Seen Via the Internet



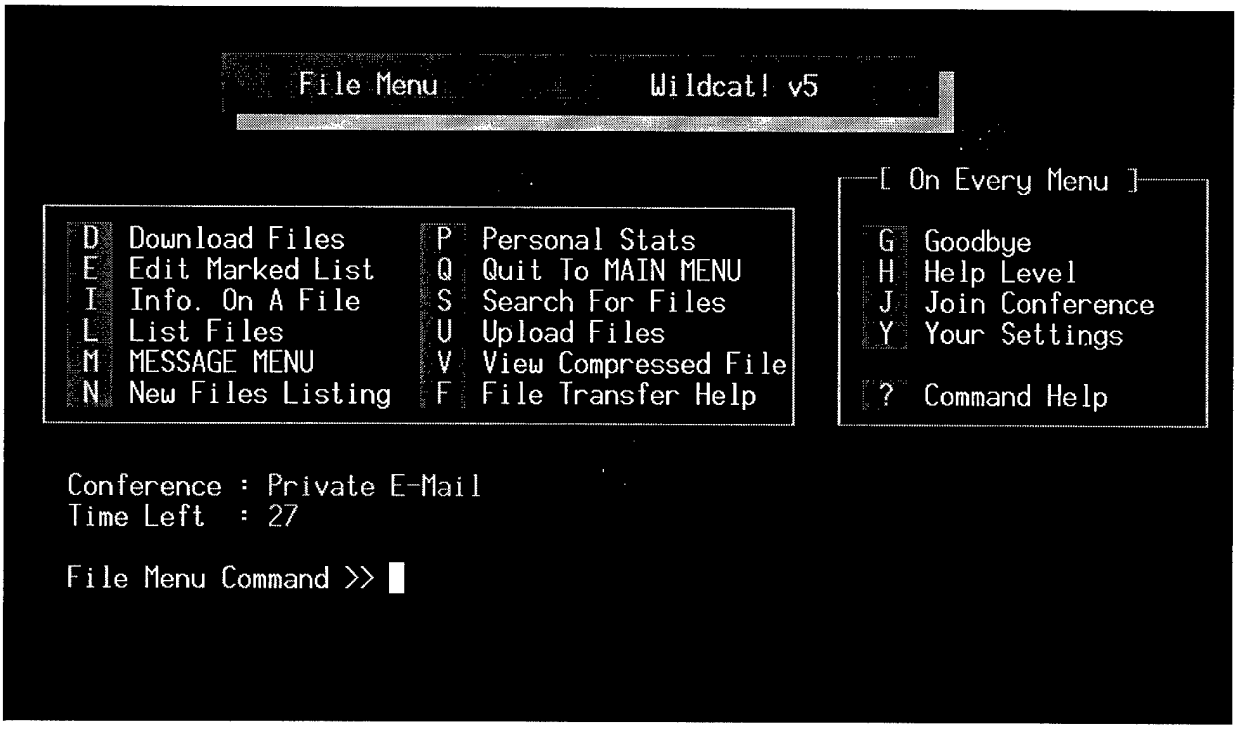
Appendix A-2
California Regional Bulletin Board Menus as Seen Via the Dial-in BBS



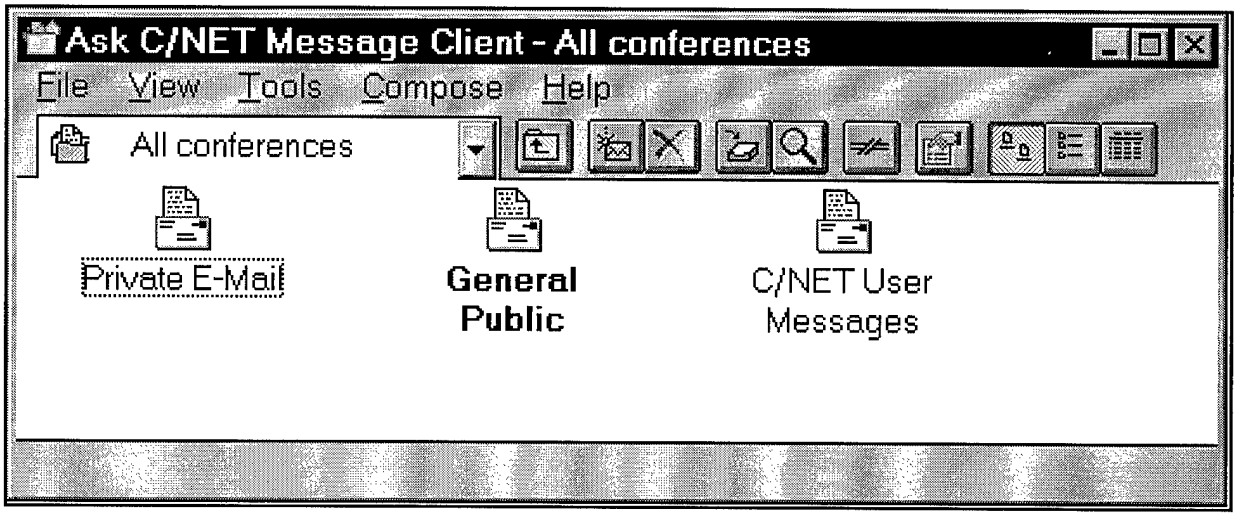
Appendix A-3
BBS File Menu as Seen via the Internet



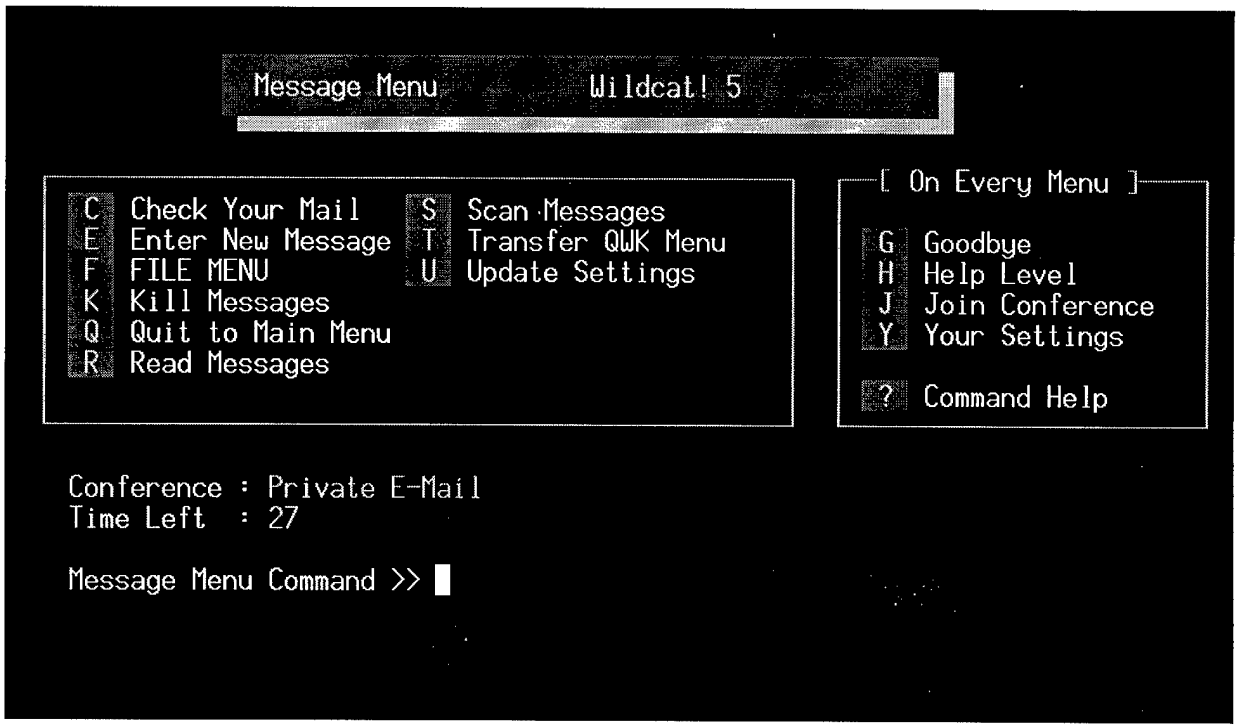
Appendix A-4: BBS File Menu as Seen via the Dial-in BBS



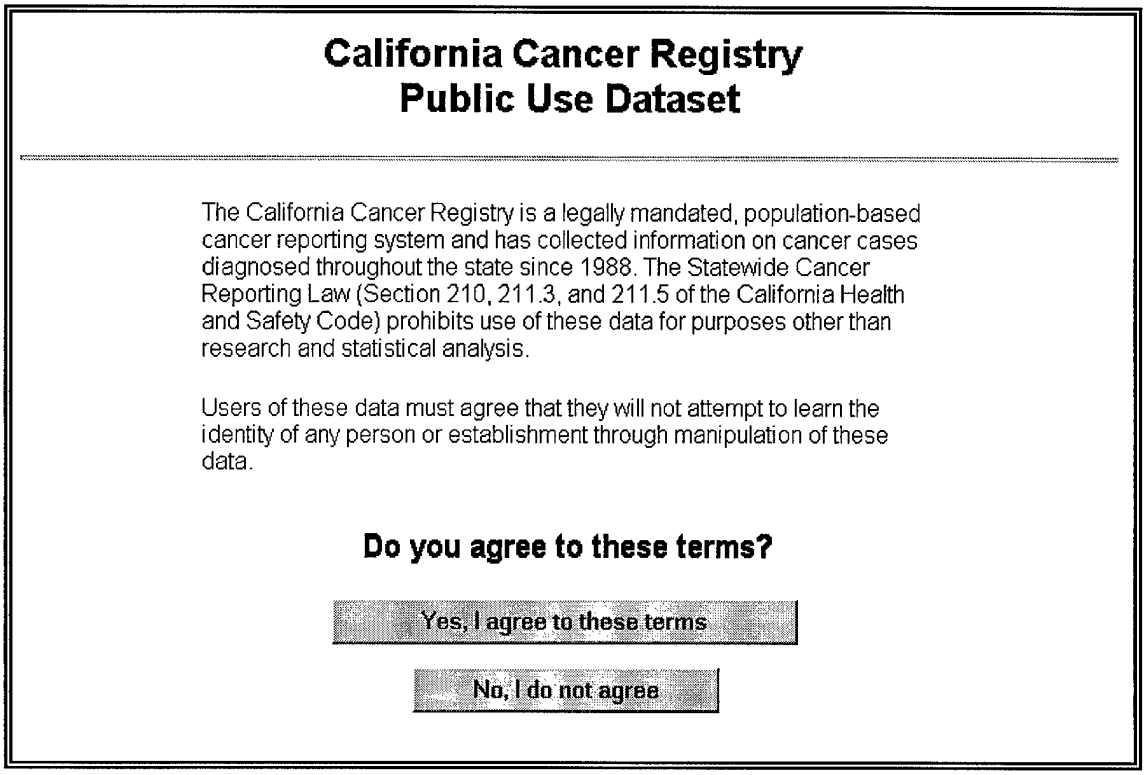
Appendix A-5
BBS Mail Menu as Seen via the Internet

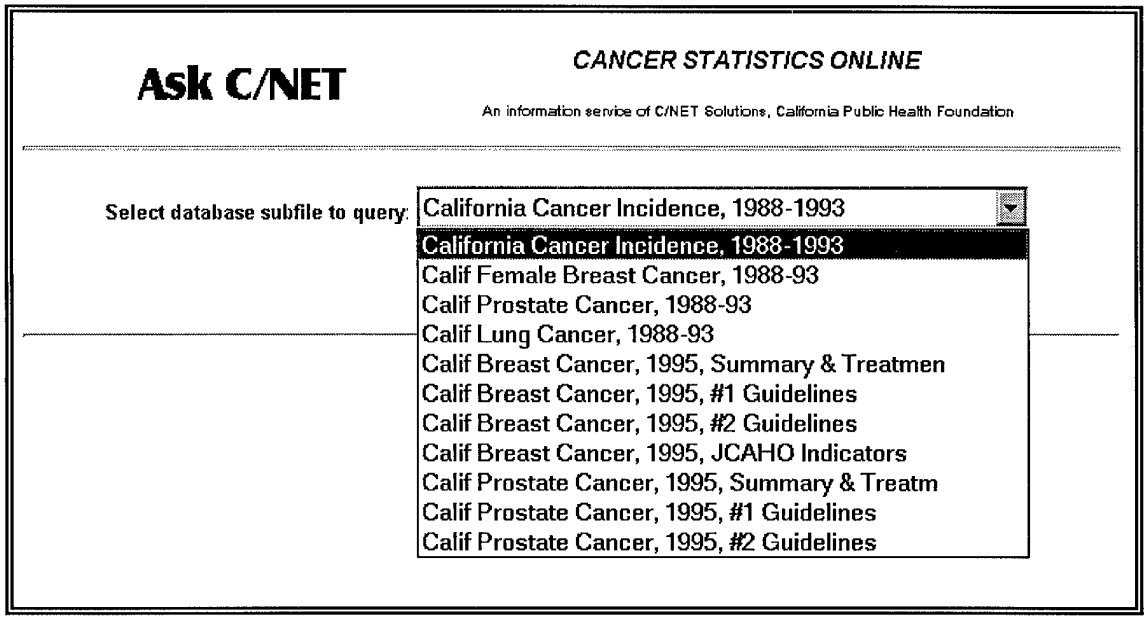


Appendix A-6
BBS Mail Menu as seen via the Dial-in BBS

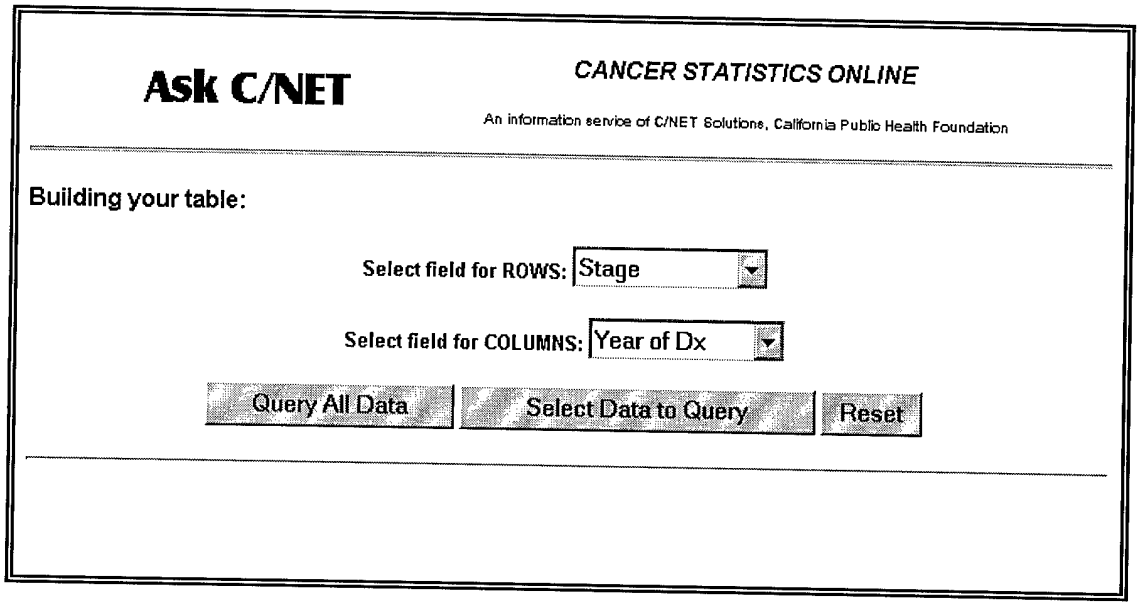


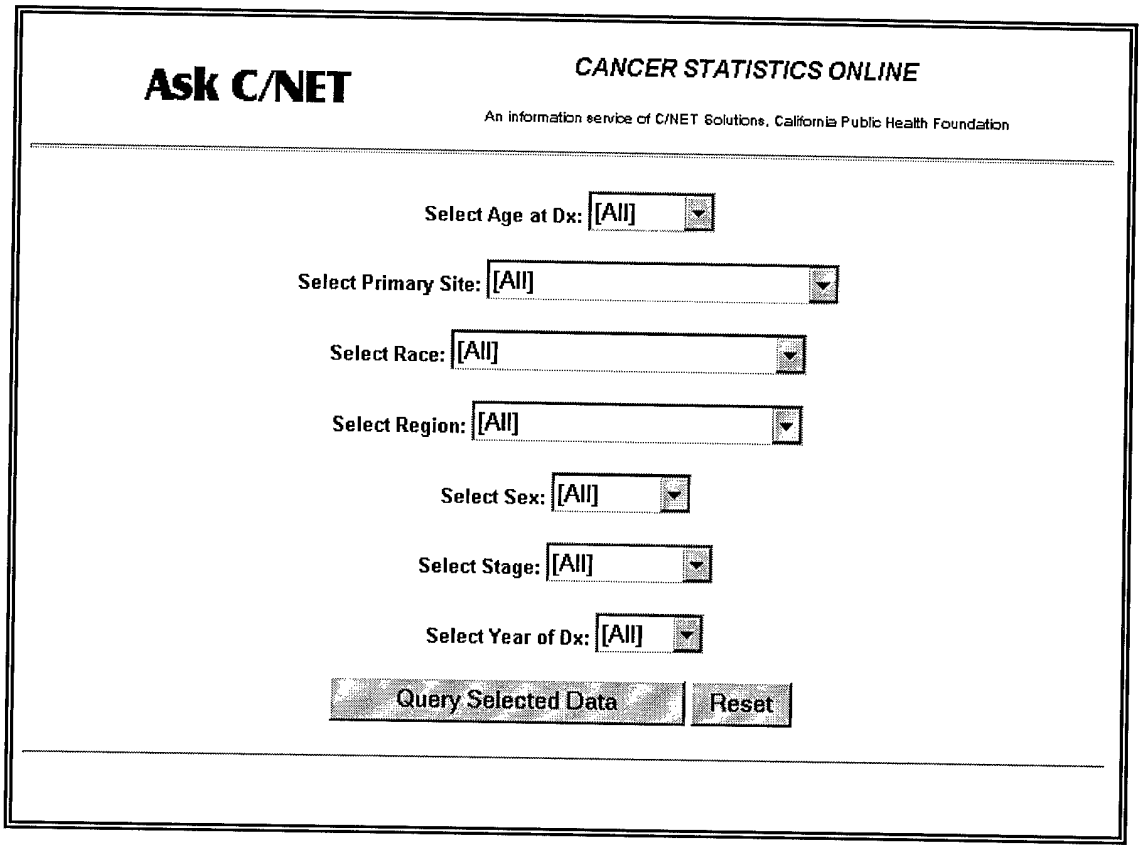
Appendix B:
Accessing the California Cancer Registry Public Use File, 1988-1993





Appendix B (Continued)





Appendix B (Continued)

Ask C/NET

CANCER STATISTICS ONLINE

An information service of C/NET Solutions, California Public Health Foundation

Suggested Citation

Any person or group using this data shall include the following citations in any publication, presentation, or printed report using data from the State-funded cancer reporting program and the Ask C/NET query engine:

1988-1993 Public Use File	California Cancer Registry public use tape, 1988-1993. Sacramento (CA): California Department of Health Services, Cancer Surveillance Section; 1996 March. Available from: *http://www.askcnet.org/dataq/index.htm*
1995 Early Reporting File	California Cancer Registry preliminary data, 1995. Sacramento (CA): California Department of Health Services, Cancer Surveillance Section; 1996 September. Available from: *http://www.askcnet.org/dataq/index.htm*
Program	Cancer statistics on demand. Berkeley (CA): California Public Health Foundation, C/NET Solutions; 1996 August. Available from: *http://www.askcnet.org/dataq/index.htm*

NOTE: Because of the influence on cancer incidence of the age and the size of the population at risk, these case counts cannot be used to assess cancer risk, but may be helpful in assessing the public health impacts of cancer.

The following table represents cases of cancer diagnosed during the indicated time period. Unless the file was subdivided by race or stage, the data represent all racial and ethnic groups and both invasive and *in situ* cases, respectively.

California Cancer Incidence, 1988-1993

		Year of Dx					
Stage	**Total**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**
In situ	82897	11970	12364	14035	14202	15301	15025
Localized	283088	42576	42829	45506	49167	52354	50656
Regional	155479	24758	24773	25778	26437	27381	26352
Distant	161815	25221	26116	27273	28319	27927	26959
Blank	11	1	0	0	2	2	6
Unstaged	97514	16171	16024	16716	15990	16103	16510
Total	780804	120697	122106	129308	134117	139068	135508

- Display information about the California database.
- Perform another query using the California database.
- Perform a query using a different database.
- Return to the Ask C/NET Home Page.

Appendix C: Sample Output from the SEER Public Use File, 1973-1992

Ask C/NET

CANCER STATISTICS ONLINE

An information service of C/NET Solutions, California Public Health Foundation

Suggested Citation

Any person or group using this data shall include the following citations in any publication, presentation, or printed report using data from the SEER Program and the Ask C/NET query engine:

SEER Public Use File	Surveillance, Epidemiology, and End Results (SEER) Program public use CD-ROM (1973-92), National Cancer Institute, DCPC, Surveillance Program, Cancer Statistics Branch, July 1995.
Program	Cancer statistics on demand. Berkeley (CA): California Public Health Foundation, C/NET Solutions; 1996 August. Available from: *http://www.askcnet.org/dataq/index.htm*

NOTE: Because of the influence on cancer incidence of the age and the size of the population at risk, these case counts cannot be used to assess cancer risk, but may be helpful in assessing the public health impacts of cancer.

SEER Breast Cancer, 1973-92

		Radiation		
Surgery	Total	No radiation	Radiation given	Radiation unknown
No Surgery	4068	3096	902	70
Partial mastectomy no nodes	19518	13159	6127	232
Partial mastectomy with nodes	26602	4515	21434	653
Subcutaneous mastectomy	781	699	76	6
Total mastectomy no nodes	7226	6563	561	102
Total mastectomy with nodes	97654	88158	8496	1000
Radical mastectomy	723	570	137	16
Extended radical mastectomy	20	12	7	1
Regional/distant surgery	396	183	199	14
Surgery NOS	88869	69443	17928	1498
Surgery Unknown	13816	6998	3731	3087
Total	259673	193396	59598	6679

- Display information about the SEER database.
- Perform another query using the SEER database.
- Perform a query using a different database.
- Return to the Ask C/NET Home Page.

Appendix D: Accessing the SEER Public Use File, 1973-1992

SEER Public Use File

There are specific laws which insure the confidentiality of individuals diagnosed with cancer when information about their cancer is entered into a data base for the purpose of establishing a research resource. In utilizing data on such individuals for research purposes, it is absolutely necessary to insure, to the extent possible, that uses of any such data will be limited to research and that uses for any other reason, particularly those resulting in personal disclosure, will be prosecuted to the full extent of the law.

In order for the Surveillance, Epidemiology, and End Results Program to provide a public use data file to you, it is necessary that you agree to the following provisions:

1. You will not use nor permit others to use the data in any way other than for statistical reporting and analysis.
2. You will not attempt to link nor permit others to link the data with individually identified records in another data base.
3. No one having access to the data will attempt to learn the identity of any person whose cancer data is in the data base.
4. If the identity of any person is discovered inadvertantly, then the following should be done:
   - no use will be made of this knowledge,
   - the Cancer Statistics Branch will be notified of the incident,
   - no one else will be informed of the discovered identity.
5. You will not release nor permit others to release the datasets or any part of them to any person who is not a member of your organization except with the written approval of NCI.

Clicking the **"I agree"** button below indicates that you agree to comply with the above provisions. Deliberately making a false statement regarding any matter within the jurisdiction of any department or agency of the Federal Government violates 18 USC 1001 and is punishable by a fine up to $10,000 or up to five years in prison.

Do you agree to these terms?

Yes, I agree to these terms

No, I do not agree

Appendix E:

Sample Askcnet Query Log

Dateset	Row	Column	Selection
California Cancer Incidence, 1988-199	**Primary Site**	**Year of Dx**	Sex : Female Race : White, N Stage : Localize Primary Site : Cervix u Age at Dx : 30-34
Date: 8/19/96 18:23:41 **User Type: Provider**			
Calif Lung Cancer, 1988-93	**Sex**	**Year of Dx**	
Date: 8/20/96 15:22:31 **User Type: Unknown**			
Calif Lung Cancer, 1988-93	**Histology**	**Reason No Surger**	
Date: 8/20/96 15:53:46 **User Type: Unknown**			
California Cancer Incidence, 1988-199	**Primary Site**	**Year of Dx**	Sex : Female Stage : In situ Region : San Fran Primary Site : Breast
Date: 8/20/96 21:25:54 **User Type: Registrar**			
Calif Prostate Cancer, 1988-93	**Year of Dx**	**Stage**	
Date: 8/20/96 23:38:52 **User Type: Researcher**			

Appendix F: New Data Items, CCR 1995 Early Reporting File

New Data Items for 1995 Cases

JCAHO Clinical Indicators (Breast, Lung, & Colorectal)	Treatment (All Sites)	Global Clinical Guidelines (All Sites)	Other
T &N Staging Documented	Surgical Approach	Surgical Consult	Discovd by Screening
Margin Status Documented	Reason for No Radiation	Radiation Oncology Consult	Family History of Cancer
Histology Documented	Reason for No Chemotherapy	Medical Oncology Consult	Personal History of Cancer
Extension Documented	Reason for No Hormone Therapy	CBC	Support Services
Tumor Size Documented	Chemotherapy Completion Status	Multichannel Chemistry	S-Phase (Breast)
Lymph Node Examination Documented	Protocol Eligibility	Chest X-Ray	DNA Level (Breast)
Surgical Path Consult Documented		MRI of Primary	PSA Value (Prostate)
Staged by Managing Physician		CT Chest/Lung	PSA Level (Prostate)
ER Analysis Documented (Breast)		CT Abdomen/Pelvis	
Complete Resection (Lung)		CT Liver/Spleen	
Barium Enema (Colorectal)		Imaging Bone	
Colonoscopy (Colorectal)		Imaging Brain	
Obstructed or Perforated (Colorectal)			
Proctosigmoidoscopy (Colorectal)			

Appendix G: California Cancer Registry, 1995 Early Reporting

Ask C/NET

CANCER STATISTICS ONLINE

An information service of C/NET Solutions, California Public Health Foundation

Suggested Citation

Any person or group using this data shall include the following citations in any publication, presentation, or printed report using data from the State-funded cancer reporting program and the Ask C/NET query engine:

1988-1993 Public Use File	California Cancer Registry public use tape, 1988-1993. Sacramento (CA): California Department of Health Services, Cancer Surveillance Section; 1996 March. Available from: *http://www.askcnet.org/dataq/index.htm*
1995 Early Reporting File	California Cancer Registry preliminary data, 1995. Sacramento (CA): California Department of Health Services, Cancer Surveillance Section; 1996 September. Available from: *http://www.askcnet.org/dataq/index.htm*
Program	Cancer statistics on demand. Berkeley (CA): California Public Health Foundation, C/NET Solutions; 1996 August. Available from: *http://www.askcnet.org/dataq/index.htm*

NOTE: Because of the influence on cancer incidence of the age and the size of the population at risk, these case counts cannot be used to assess cancer risk, but may be helpful in assessing the public health impacts of cancer.

Calif Breast Cancer, 1995 - Clin Indicators

		Stage					
T & N in path rpt	Total	*in situ*	Localized	Regional	Distant	Not staged	Unknown stage
DNA	2735	325	1294	819	98	18	181
Req met	707	66	431	196	12	0	2
Req not met	752	109	432	193	17	1	0
Justifiable exception	42	11	17	11	3	0	0
Unknown	57	5	25	21	3	0	3
Total	4293	516	2199	1240	133	19	186

- Display information about the California database.
- Perform another query using the California database.
- Perform a query using a different database.
- Return to the Ask C/NET Home Page.

Appendix H: Rules for Flagging Missing Breast Cancer Treatment

Adapted from Fred Hutchinson Cancer Research Center, 1996, based on the NCI PDQ system

Recommended Treatment is

If Stage is *in situ* [*i.e.*, (Stage = 0, Tis, N0, M0) or Behavior = 2]
Then if Histology is 85202 (lobular *in situ*)
Then Definitive Surgery would be done (code 10-98)
AND
Hormone Therapy would be given (code 1-6)

Else [*i.e.*, Histology is not 85202]
Then Definitive Surgery would be done (code 10-98)
AND
Hormone Therapy would be given (code 1-6)
AND
Radiation Therapy would be given (code 1-6)

Else [*i.e.*, Stage is not *in situ*]
Then Definitive Surgery would be done (code 10-98)
AND
Hormone Therapy would be given (code 1-6)
AND
Radiation Therapy would be given (code 1-6)
AND
Chemotherapy would be given (code 1-6)

Appendix I: Clinical Trial Matching System Profile Query Web Form

Clinical Trial Matching System

Your Profile

After you complete the questionnaire below, the CTMS will match the information to its database of current clinical trials to see if there are any trials for which you may be eligible. Please try to be as accurate as you can in answering the questions so that the system can make the best match possible.

What is your age?

When were you diagnosed with breast cancer? (mm/dd/yy)

What stage disease were you diagnosed with?

0
0
1
2A
2B
3A
3B
4

What kind of treatment have you received already?

(to check more than one type of treatment, hold the **control key** down while you click on each type)

(for explanation of treatment terms, check the Glossary)

- ...only
- ...tomy
- ...Node Dissection
- ...aneous Mastectomy
- Simple Mastectomy
- Total Mastectomy
- Modified Radical Mastectomy
- Radical Mastectomy
- Extended Radical Mastectomy
- Mastectomy
- Some kind of Surgery
- Chemotherapy
- Hormone Therapy
- Bone Marrow Transplant

If you have had surgery, when did you have it? (mm/dd/yy)

Do any of these terms describe you?

(to check more than one description, hold the control key down while you click on each one)

I have cancer in both breasts

(for explanation of terms, check the Glossary)

MATCH ME | CLEAR MY ANSWERS

High Dose Chemotherapy and Bone Marrow Transplant vs. Intensive Dose Chemotherapy in Stage II - IIIA Breast Cancer with ≥ 10 Nodes

Appendix J: Example of Clinical Trial Description

CTIP RANDOMIZED

STAGE II-IIIA (regional disease)
Phase III
Financed by NCI

WHO IS IT FOR: Women 18 or older in all menopausal stages
*** This study requires hospitalization**

Purpose of this Study

Identify whether bone marrow transplant is more effective than high dose chemotherapy by comparing the side effects, as well as disease-free and overall survival of women that have Stage 2 or 3 breast cancer with 10 or more positive lymph nodes.

Official Title

Phase III Randomized Comparison of High-Dose Cyclophosphamide/Cisplatin/Carmustine (CTX/DCCP/BCNU) with Autologous Marrow and Peripheral Stem Cell Support vs. Standard-Dose CTX/CDDP/BCNU Following Adjuvant Cyclophosphamide/Doxirubicin/ Fluorouracil (CAF) in Women with Stage II/IIIA Breast Cancer with at Least 10 Positive Axillary Nodes (CALGB 9082, SWOG 9114)

About the Study

All patients receive CAF chemotherapy before they are placed in treatment "Arm" I or II. If the woman is assigned to Arm I, she will receive High Dose chemotherapy (CTX/DCCP/BCNU), a Bone Marrow and Stem Cell transplant with G-CSF (to help stimulate her blood cells for recovery), Radiation, and Tamoxifen if her hormone receptors are ER+. If the woman is assigned to Arm II, she will receive an Intensive Dose chemotherapy (CTX, CDDP, BCNU), G-CSF (if needed), radiation, and Tamoxifen if ER+. Results will be studied by institution, disease stage, and type of ER and PR. Patients will be followed closely for 2-1/2 years, and annually after that.

☐ *Required*	♋ *Permitted*	☒ *Not Allowed*
• Documented insurance coverage for this trial	• All menopausal stages ok (before, after and during)	• Other cancers (cervical or non-melanoma skin cancer ok)
• Stage II or III breast cancer	• All types of ER and PR receptors	• Breast cancer in both breasts
• 10 or more positive lymph nodes		• Heart disease
• Mastectomy or lumpectomy within 8 weeks of starting CAF		• Consistent numbness or hearing loss
• Tissue around tumor free of disease		• Prior chemotherapy
• Normal liver, kidney, heart and lungs		• Radiotherapy before this treatment

Please see "Drugs and their Effects" and "Glossary" sections for more information and definitions.
Clinical trial information can change quickly. Please discuss this with medical professionals to receive up to date information.

High Dose Chemotherapy and Bone Marrow Transplant vs. Intensive Dose Chemotherapy in Stage II - IIIA Breast Cancer with ≥ 10 Nodes

- No blood viruses (HIV, Hepatitis C)
- Normal blood counts
- No identified metastasis
- Must be over 18 years old
- Medical or mental illnesses
- Pregnancy or breastfeeding

Who to Call for More Information

City/Area	Trial Location	Contact Person	Phone
Concord	Mount Diablo Medical Center	Michael Messer	XXXXXXXXXXXX
Daly City	Seton Medical Center	John Siebel, M.D.	XXXXXXXXXXXX
Oakland		Karen Egan	XXXXXXXXXXXX
Santa Rosa	Northern CA Cancer Center	Sue Silkworth	XXXXXXXXXXXX
Sacramento	UC Davis	Frederick Meyer, M.D.	XXXXXXXXXXXX
San Francisco	UCSF	Nick Jorgenson	XXXXXXXXXXXX

Known Side Effects

Due to the high doses of chemotherapy in both arms of this study, there can be serious side effects involving the heart, lungs, skin, stomach, liver, kidney, bladder, blood, nervous system, and immune system. Rarely, there are additional illnesses and cancers that could also occur. Approximately 2-15% of patients die as a result of this kind of treatment. Chemotherapy also causes menopause in most younger women.